PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 15/29, C07K 14/415, C12N 5/10, A01H 5/00 | A2

(11) International Publication Number: **WO 98/08961**

(43) International Publication Date: 5 March 1998 (05.03.98)

(21) International Application Number: PCT/IB97/01037

(22) International Filing Date: 28 August 1997 (28.08.97)

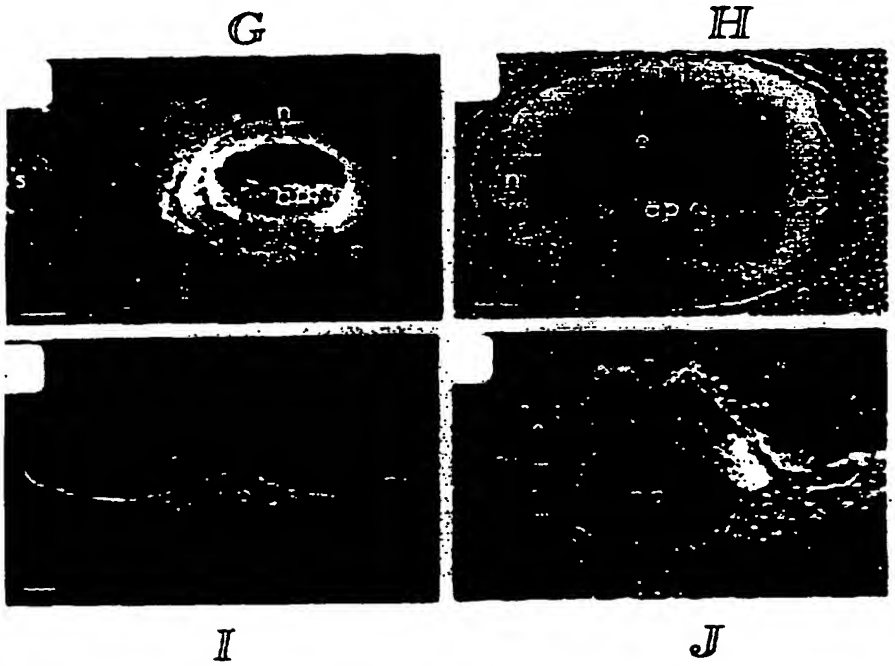
(30) Priority Data:
60/024,886 30 August 1996 (30.08.96) US

(71)(72) Applicants and Inventors: OLSEN, Odd-Arne [NO/NO]; Taarnveien 16, N-1430 Aas (NO). DOAN, Danny, N., P. [AU/NO]; Lyngveien 14B, N-1430 Aas (NO). LINNESTAD, Casper [NO/NO]; Ovre Linnestad, N-1540 Vestby (NO).

(81) Designated States: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: ENDOSPERM AND NUCELLUS SPECIFIC GENES, PROMOTERS AND USES THEREOF



(57) Abstract

Isolated nucleic acid molecules are provided which encode END1 and NUC1, endosperm proteins. Also provided are vectors which are capable of expressing such nucleic acid molecules, host cells which contain such vectors, and polypeptides encoded by the aforementioned nucleic acids. In addition, nucleic acid molecules are provided which comprise END1 or NUC1 promoters.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Description

ENDOSPERM AND NUCELLUS SPECIFIC GENES, PROMOTERS AND USES THEREOF

Technical Field

The present invention relates generally to endosperm specific genes in plants, and more specifically, to coding regions and promoters of these genes.

Background of the Invention

In cereal grain plants, such as barley, the endosperm is the main storage organ, providing nutrients to the developing and to the germinating embryo. Altering the types or amounts of nutrients in the endosperm could improve cereal grains as a food source for the world. In addition, developing seeds are susceptible to diseases caused by pathogens and insects. Increasing plant defenses at this stage of development could improve crop yields.

Currently, genetic engineering of plants involves transformation of the entire plant and generally does not allow control of expression of transgenic products in an organ-specific manner. In particular, elements required for expression of products early in embryo development are not available.

In view of the problems in obtaining expression of transgenic products in developing embryos of plants, there is a compelling need for improved expression. The present invention exploits the use of plant genes that are highly and specifically expressed in developing grains to alter the storage nutrients and capabilities of disease defenses, while providing other related advantages.

Summary of the Invention

Within one aspect of the present invention, an isolated nucleic acid molecule comprising an END1 promoter is provided. Within a related aspect, a NUC1 promoter is provided. In yet other aspects, isolated nucleic acid molecules encoding a product of END1 and NUC1 are provided.

In another aspect, a vector comprising one of the nucleic acid molecules is provided. In preferred embodiments, the vector is an expression vector. When the nucleic acid molecule comprises the END1 or NUC1 promoter, the vector further comprises a foreign gene operably linked to the promoter. In a preferred embodiment, the foreign gene confers disease resistance.

In other aspects, methods of producing a foreign gene product or a plant that expresses a foreign gene are provided.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, various references are set forth below which describe in more detail certain procedures or compositions (*e.g.*, plasmids, etc.), and are therefore incorporated by reference in their entirety.

Brief Description of the Drawings

Figure 1 shows light micrograph depicting the morphology and anatomy of barley ovaries and young grains. A. Light micrograph of intact ovary at 5 DAP. B. Light micrograph of transverse section (2 μm) through the mid-portion of 5 DAP ovary. From the outside and in, the pericarp consists of an epidermal layer, several layers of parenchyma cells, two layers of photosynthetic cross cells (c) and an inner epidermis of tube cells (the crushed cell layer) (cc). Testa (t) has an outer epidermis of thin walled cells and an inner layer of thick walled cells, the nucellus (n) has an outer epidermis and several layers of inner cells which at 5 DAP are present only on the ventral side. In the center of the grain at the ventral side, nucellus forms the nucellar projection (np), which extends into the endosperm cavity. The endosperm coenocyte (c) consists of one layer of NCDs at the arboreal stage surrounding the large central vacuoic (cv). C. Right half of transverse section of 9 DAP grain showing young cellular endosperm with anticlinal walls (a) extending to the center of the endosperm from both sides and periclinal walls (p) in the periphery. The nucellar epidermis (ne) consists of highly vacuolated cells on the outside of the endosperm. D. Scanning electron micrograph of isolated embryosac covered by testa

and nucellus. Bars represent 1 mm in A, 200 μm in B, 100 μm in C and 500 μm in D.

Figure 2 presents *in situ* hybridization analyses of END1 and NUC1 transcripts in the barley endosperm coenocyte and nucellar tissues. A. Dark-field micrograph showing the presence of END1 transcripts in the endosperm coenocyote over the nucellar projection at 6 DAP. B. Higher magnification of a selected area of A. C. Same as in A. The picture is a double exposure composed of a dark-field micrograph showing silver grains in yellow and a phase contrast micrograph enhancing histological details. D. Higher magnification of a selected area of C. E. Accumulation of the END1 transcripts on the ventral side of the endosperm over the nucellar projection right after endosperm cellularization at 8 DAP. F. END1 transcripts in modified aleurone cells and ventral starchy endosperm cells at 10 DAP. G. Dark-field micrograph of longitudinal section of unfertilized ovary (0 DAP) showing the presence of NUC1 transcripts in the nucellus. H. NUC1 transcripts in the nucellus of transverse section of a 2 DAP grain. The picture is a double exposure composed of a darkfield micrograph showing silver grains in yellow and a phase contrast micrograph enhancing histological details. I. NUC1 transcripts in the nucellar epidermis and lateral part of the nucellar projection in transverse section of a 7 DAP grain. J. Accumulation of NUC1 transcripts in the lateral elements of the nucellar projection and in the nucellar epidermis at 10 DAP. Bars in figures A, C, E, G and I represent 200 μm; B, D, and F, 80 μm; H and J, 100 μm. NP, nucellar projection; NE, nucellar epidermis; P, pericarp; N, nucellus; CV, central vacuole; ES, endosperm synctium; V, ventral vascular strand; AP, antipodal cells; S, stigma.

Figure 3 is a Northern blot analysis of END1 and NUC1 transcripts in developing barley grains. Poly(A)-rich RNA (100 ng per lane) of intact ovaries and grains from 0 to 30 DAP were hybridized to single stranded probes of the END1 and NUC1 cNDAs. The lower panel shows a control hybridization with a barley histone H3 probe. The same blot was used for all three probes.

Figure 4 presents nucleotide sequences of the END1 and NUC1 cDNA clones isolated in the differential screening experiment. These sequences were used as probes in the *in situ* and Northern blot hybridizations of Fig. 2 and 3.

Detailed Description of the Invention

Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms that will be used hereinafter.

As used herein, "END1" and "NUC1" refer to genes that are preferentially expressed in endosperm and nucellus, respectively. The amino acid sequences of representative END1 and NUC1 gene products from barley (*Hordeum vulgare* L. cv. Bomi) have been deduced and are presented in SEQ ID NOS:3-6. Within the context of this invention, it should be understood that END1 and NUC1 include wild-type protein sequence, as well as other variants (including alleles) of the native protein sequence. Briefly, such variants may result from natural polymorphisms or be synthesized by recombinant methodology, and differ from wild-type protein by one or more amino acid substitutions, insertions, deletions, or the like. Typically, amino acid substitutions are conservative. In the region of homology to the native sequence, variants should preferably have at least 90% amino acid sequence identity, and within certain embodiments, greater than 92%, 95%, or 97% identity. As will be appreciated by those skilled in the art, a nucleotide sequence encoding END1, NUC1 or variant may differ from the native sequences presented in SEQ ID NOs:1 and 2, due to codon degeneracies, nucleotide polymorphisms, or amino acid differences.

As used herein, a "promoter" refers to a nucleotide sequence that contains elements that direct the transcription of a linked gene. At minimum, a promoter contains an RNA polymerase binding site. More typically, in eukaryotes promoter sequences contain binding sites for other transcriptional factors that control the rate and timing of gene expression. Such sites include TATA box, CAAT box, POU box, AP1 binding site, and the like. Promoter regions may also contain enhancer elements.

An "isolated nucleic acid molecule" refers to a polynucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid construct, that has been separated from its source cell (including the chromosome it normally resides in) at least once in a substantially pure form. Nucleic acid molecules may be comprised of a wide variety of nucleotides, including DNA, RNA, nucleic acid analogues, or some combination of these.

A nucleotide molecule having "END1 promoter activity" refers to a promoter region containing elements responsive to factors that control END1 gene expression. In general, such a sequence promotes a similar expression pattern as for native END1 (*e.g.*, tissue specificity, developmental timing). A sequence with END1 promoter activity may vary from the native sequence by base substitutions, insertions, and deletions, for example. Such alterations may affect the relative strength of the promoter, but should not affect the expression pattern. A nucleotide sequence having "NUC1 promoter activity" is defined in similar fashion as for a sequence having END1 promoter activity.

A. END1 AND NUC1 GENES AND GENE PRODUCTS

As noted above, the present invention provides compositions relating to END1 and NUC1 genes, and methods for the use of the END1 and NUC1 gene products.

Given the disclosure provided herein, the END1 and NUC1 genes, as well as other developmental-specific genes, including endosperm-specific and nucellus-specific genes, can be readily isolated from a variety of plants, including for example, barley (*Hordeum*). For example, within one embodiment, genes encoding END1 and NUC1 can be identified by differential hybridization. Briefly, in this method, a cDNA library is constructed, such as in a lambda vector. The recombinant clones are plated and clone DNA is transferred to a solid support (*e.g.*, nylon filters). Labeled probe is prepared from RNA isolated from a positive source, such as endosperm, and a negative source, such as pericarp. Probe is hybridized to duplicate filters and clones hybridizing with the positive, but not with negative, probe are identified and isolated.

Other methods for identifying differentially expressed genes are also readily available (*see, for example,* Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989; Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing, 1987). Briefly, the methods generally fall into three major categories: subtractive hybridization, differential screening, and random-primed amplification. Combinations and variations of these methods are also widely used. For example, a library pre-enriched for differentially expressed transcripts by subtactive hybidization can be prepared (*see, e.g.,* Sive and St. John, *Nucl. Acids Res. 16*: 10937, 1988) and screened as above. Random-primed PCR is also widely employed (*see, e.g., Methods Enzymol. 254*: 275, 1995; *Trends Genet. 11*: 242, 1995; Liang and Pardee, *Science 257*: 967, 1992; Welsh et al., *Nucl. Acids Res. 20*: 4965, 1992). In addition, variations of random-primed PCR may also be used, especially when a particular gene or gene family is desired. In such a method, one of the amplification primers is an anchored oligo(dT) (oligo(dT)dN) and the other of the primers is a degenerate primer based upon amino acid or nucleotide sequence of a related gene.

Alternatively, other methods may be used to obtain a nucleic acid molecule that encodes END1, NUC1, or other developmental-specific gene. For example, a nucleic acid molecule may be obtained from a cDNA or genomic expression library by screening with an antibody or antibodies reactive to one or more of these gene products (*see*, Sambrook, et al. *supra;* Ausubel, et al. *supra*). Or, using the sequence information provided herein, a probe sequence can be synthesized and labeled, such as with a radioactive label, enzymatic label, protein label, fluorescent label, or the like, and hybridized to a genomic library or a cDNA library constructed in a phage, plasmid, phagemid, or other viral vector (*see, for example,* Sambrook et al., *supra*, Ausabel, et al., *supra*). DNA representing RNA or genomic nucleic acid sequence may also be obtained by amplification using sets of primers complementary to 5' and 3' sequences of the cDNA sequence, such as presented in SEQ ID NOs:1 and 3. For ease of cloning, restriction sites may also be incorporated into the primers.

Variants (including alleles) of the END1 and NUC1 gene products provided herein may be readily isolated from natural variants (*e.g.*, polymorphisms, mutants), synthesized or constructed. Many methods have been developed for generating mutants (*see, generally*, Sambrook et al., *supra*; Ausabel, et al., *supra*). Briefly, preferred methods for generating a few nucleotide substitutions utilize an oligonucleotide that spans the base or bases to be mutated and contains the mutated base or bases. The oligonucleotide is hybridized to complementary single stranded nucleic acid and second strand synthesis is primed from the oligonucleotide. The double-stranded nucleic acid is prepared for transformation into host cells, typically *E. coli*, but alternatively, other prokaryotes, yeast or other eukaryotes. Standard screening and vector growth protocols are used to identify mutant sequences and obtain high yields.

Similarly, deletions and/or insertions of the END1 and NUC1 genes may be constructed by any of a variety of known methods. For example, the gene can be digested with restriction enzymes and religated such that sequence is deleted or religated with additional sequence such that an insertion or large substitution is made. Other means to generate variant sequences may be found, for example in Sambrook et al. (*supra*) and Ausubel et al. (*supra*). Verification of variant sequences is typically accomplished by restriction enzyme mapping, sequence analysis, or probe hybridization.

B. PROMOTER REGIONS OF END1 AND NUC1 GENES

As described above, this invention provides promoters of END1 and NUC1 genes and uses of the promoters in controlling foreign gene expression.

Briefly, the promoter regions of END1 and NUC1 genes are initially identified in a genomic clone by DNA sequence analysis. The sequence from a cDNA is located in the genomic clone by direct DNA sequence analysis, restriction mapping, RNase protection, S1 nuclease analysis, or other similar methods. The DNA sequence upstream of the transcribed region is determined. Because promoters are relatively close to transcribed sequences, identification of a promoter may be made by DNA sequence analysis and/or by functional analysis.

Identification of classical consensus promoter sequences, such as a TATA box and CAAT box, located adjacent to and upstream of the coding region indicate the presence of a promoter. In addition, other regulatory elements in the promoter region, such as an enhancer, wound inducible factor sequences and the like, that have identifiable sequences may be present. Verification of a promoter by functional analysis is assessed by cloning the putative promoter sequence into a plasmid containing a reporter gene, transfecting a cell capable of utilizing the promoter, and detecting expression of the reporter gene. Briefly, a DNA fragment containing the candidate promoter sequence is inserted into a vector containing a reporter gene, such as pBI121, which contains GUS (β-glucuronidase gene) under control of the CaMV 35S promoter (Jefferson et al., *EMBO J. 6*: 3901-3907, 1987), in place of the promoter. If a promoter is present on the inserted fragment, transcription of GUS will ensue in cells that can activate the promoter. GUS activity is detected by enzymatic activity on any one of the numerous, readily assayable GUS substrates. One convenient substrate is the colorometric substrate 5-bromo-4-chloro-3-indoyl-glucuronide sodium salt. Other substrates that give a colorometric or fluorogenic signal may be readily obtained and used.

This method may be generalized to assay promoter activity of any candidate promoter sequence on a DNA fragment by constructing a vector containing the DNA fragment upstream of a suitable reporter gene. Reporter genes include, but are not limited to, GUS, luciferase, β-galactosidase, and green fluorescent protein. Preferably, the reporter gene has no endogenous counterpart in the host cell, or any endogenous gene is expressed minimally or not at all. Examples of suitable constructs are discussed below.

The END1 or NUC1 promoter in this construct may be tested directly in plants, preferably in cereal plants. Briefly, for gene promoters described herein, the construct described above is used to transform plants, such as tobacco leaf discs, via *Agrobacterium tumefaciens*-mediated transformation. GUS activity is then detected in tissue sections of transformed plants. By assaying different plant tissues, a pattern of expression is established. Typical plant tissues for assay include floral

regions of the plant, such as developing and germinating seed, vasculature, photosynthetic tissues, such as mature leaves, and stems. Alternatively, promoter activity may be tested in a transient expression assay in plants. In such a method, a construct containing the candidate promoter region is placed upstream of a reporter gene (*e.g.*, GUS, luciferase, β-galactosidase, green fluorescent protein) and transfected into plant cells by bombardment (*see, GUS Protocols: Using the GUS gene as a reporter of gene expression*, Gallagher (ed.), Academic Press, Inc., 1992). Reporter activity is measured a few hours later. When using this method, the plasmid or DNA may be an *Agrobacterium*-based plasmid, pUC-based plasmid, or other vector. Various plant tissues may be used, which will provide some information regarding cell-specificity of the promoter. Approximately 100 mg of tissue is needed for bombardment.

Minimal promoter sequences and variant promoter sequences are constructed by standard techniques, such as deletion by restriction digestion or other nucleolytic digestion techniques. By making progressively larger deletions from the 5' end of the initial fragment and assaying each deletion mutant for promoter activity, the minimal promoter region is determined. In general, promoter activity will be found in the first 200 to 500 bp upstream of the transcription initiation site. In some cases, the promoter may extend further or less far.

C. FOREIGN GENES

As discussed above, this invention provides vectors for the expression of foreign or heterologous genes under control of endosperm-specific or nucellus-specific promoters, such as the END1 and NUC1 promoter regions. Within the context of this invention, a foreign gene is any gene sequence other END1 or NUC1, including for example, other proteins, antisense sequences, or ribozyme sequences.

The foreign genes are under control of an END1 or NUC1 promoter. END1 is expressed in cells at the gateway to the endosperm as early as 5 dap (days after pollination). NUC1 is expressed in the nucellus. As such, the promoters of these genes may be used to alter the flux of various storage proteins or engineer disease resistance. For example, expression of storage polysaccharides prior to starch

deposition, specific proteins, enzymes, or molecules that act as sinks for other proteins may be manipulated to occur earlier in development than normal. This disease resistance may be engendered by using the END1 and/or NUC1 promoter to drive the expression of (1) enzymes that degrade pathogen-produced molecules (*e.g.*, HC-toxin reductase, fumonis in reductase); (2) enzymes that suppress pathogen growth (*e.g.*, viral coat proteins, chitinase); (3) enzymes that produce substances that suppress pathogen growth (*e.g.*, glucose oxidase); (4) toxin pumps that render a plant insensitive to pathogen-produced molecules (*e.g.*, cercosporin pump, HC-toxin pump, DON pump); and (5) proteins that activate plant defenses (*e.g.*, bacterial opsin). The choice of the foreign gene depends in part upon the desired result. For example, when disease or insect resistance to a pest or pathogen, a preferred gene is specific to the disease or insect.

As noted above, the present invention provides vectors capable of expressing genes under the control of the END1 or NUC1 promoter. In general, the vectors should be functional in plant cells, although at times, it may be preferable to have vectors that are functional in *E. coli* or other organisms (*e.g.*, production of protein for raising antibodies, DNA sequence analysis, construction of inserts, obtaining quantities of nucleic acids). Vectors and procedures for cloning and expression in *E. coli* are discussed herein and, for example, in Sambrook et al. (*supra*) and in Ausubel et al. (*supra*).

Vectors that are functional in plants are preferably binary plasmids derived from *Agrobacterium* plasmids. Such vectors are capable of transforming plant cells. These vectors contain left and right border sequences that are required for integration into the host (plant) chromosome. At minimum, between these border sequences is the gene to be expressed under control of a promoter. In preferred embodiments, a selectable marker and a reporter gene are also included. For ease of obtaining sufficient quantities of vector, a bacterial origin that allows replication in *E. coli* is preferred.

The vector should contain a promoter sequence. Preferably, for expression of a foreign gene, the promoter is an END1 or NUC1 promoter. The

entire promoter region does not need to be in the vector. However, the vector should contain at least the minimum sequence to promote transcription of the associated gene. Delineation of the minimum sequence is discussed above. In general, a minimum promoter region is contained within the first approximately 200 bases upstream of the transcription start site, but may extend for 300-500 bases. As well, variants of the promoter region may be used as long as at least 1% of native promoter activity is retained, and more preferably, greater than 10%, 20%, or 50% of native promoter activity. As provided herein, variants may be the result of natural polymorphisms, or synthesized mutants. Variants may also be derived from highly related promoter sequences, such as END1 or NUC1 promoter sequences from other plants, isolated by hybridization for example.

In certain preferred embodiments, the vector contains a reporter gene. The reporter gene should allow ready determination of transformation and expression. The GUS (β-glucoronidase) gene is preferred (U.S. Patent No. 5,268,463). Other reporter genes, such as β-galactosidase, luciferase, green fluorescent protein, and the like, are also suitable in the context of this invention. Preferably, the reporter gene does not have an endogenous counterpart. Methods and substrates for assaying expression of each of these genes are well known in the art. The reporter gene should be under control of a promoter that is functional in plants. Such promoters include CaMV 35S promoter, mannopine synthase promoter, ubiquitin promoter and DNA J promoter.

Preferably, the vector also contains a selectable marker for identifying transformants. The selectable marker may confer a growth advantage under appropriate conditions. Generally, selectable markers are drug resistance genes, such as neomycin phosphotransferase. Other drug resistance genes are known to those in the art and may be readily substituted. The selectable marker has a linked promoter, which may be constitutive or inducible, and a termination sequence, including a polyadenylation signal sequence.

A bacterial origin of replication and a selectable marker for bacteria are preferably also included in the vector. Of the various origins readily available

(*e.g.*, colEI, fd phage), a colEI origin of replication is preferred. Most preferred is the origin from the pUC plasmids, which allow high copy number.

A general vector suitable for use in the present invention is based on pBI121 (U.S. Patent No. 5,432,081) a derivative of pBIN19. Other vectors have been described (U.S. Patent No. 4,536,475) or may be constructed based on the guidelines presented herein. The plasmid pBI121 contains a bacterial origin of replication, selectable marker, and a left and right border sequence for integration into a plant host chromosome. These border sequences flank two genes. One is a kanamycin resistance gene (neomycin phosphotransferase) driven by a nopaline synthase promoter and using a nopaline synthase polyadenylation site. The second gene is the *E. coli* GUS gene under control of the CaMV 35S promoter, which is polyadenylated using a nopaline syntase polyadenylation site. The CaMV 35S promoter is excised by appropriate restriction digestion. For example, a *Hind* III/*Bam*H I double digestion or a partial *Sph* I/ complete *Xba* I digestion will excise the CaMV promoter sequence. The END1 or NUC1 promoter is then inserted in its place. These promoter sequences may be obtained by amplification, automated synthesis or isolated from a clone. Compatible cloning restriction sites are added by standard methodologies, such as the addition of adaptors or linkers. The vector is transformed into bacteria, propagated, and/or transformed into plants.

D. PLANT TRANSFORMATION METHODS

As discussed above, the present invention also provides methods for producing a plant which expresses a foreign gene, comprising the steps of (a) introducing a vector as described above into an embryogenic plant cell, wherein the vector contains a foreign gene in an expressible form, and (b) producing a plant from the embryogenic plant cell, wherein the plant expresses the foreign gene.

Vectors may be introduced into plant cells by any of several methods. For example, DNA may be introduced as a plasmid by *Agrobacterium* in co-cultivation or bombardment. Other transformation methods include electroporation, $CaPO_4$-mediated transfection, and the like. Preferably, DNA is first transfected into *Agrobacterium* and subsequently introduced into plant cells. Most preferably, the

infection is achieved by co-cultivation. In part, the choice of transformation methods depends upon the plant to be transformed. For example, monocots generally cannot be transformed by *Agrobacterium*. Thus, *Agrobacterium* transformation by co-cultivation is most appropriate for dicots and for mitotically active tissue. Non-mitotic dicot tissues can be efficiently infected by *Agrobacterium* when a projectile or bombardment method is utilized. Projectile methods are also generally used for transforming sunflowers and soybean. Bombardment is used when naked DNA, typically *Agrobacterium* or pUC-based plasmids, is used for transformation or transient expression.

Briefly, co-cultivation is performed by first transforming *Agrobacterium* by freeze-thawing (Holsters et al., *Mol. Gen. Genet. 163*: 181-187, 1978) or by other suitable methods (*see,* Ausubel, et al., *supra*; Sambrook et al., *supra*). *Agrobacterium* containing the plasmid are grown overnight at 28°C with continuous agitation in YEP medium in the presence of kanamycin (when the selectable marker is kan^R) or colony selected and purified on agar-containing medium. An aliquot or colony is grown to mid-log phase (*e.g.*, OD=0.5) in medium containing 12.5 mM MES, 1 g/L NH4Cl and 0.3 g/L $MgSO_4$ at pH 5.7. Approximately 10^9 cells/ml *Agrobacterium* is incubated with sterile leaf disks, protoplasts or meristematic tissue for 1 hr. The discs are then washed in sterile distilled water and cultivated on standard plant tissue culture medium containing kanamycin.

For microprojectile bombardment, seeds are surface sterilized in 20% bleach solution with two drops of Tween 20 per 50 ml for 30 min and rinsed twice with distilled water. Seeds are then imbibed in distilled water for 60 min, and the cotyledons are broken off to produce a clean fracture at the plane of the embryonic axis. The explants are then bisected longitudinally between the primordial leaves. The explants are placed cut surface up on GBA medium with mineral and vitamin additives, sucrose, indole-3-acetic acid, gibberellic acid and phytagar. Thirty to forty explants are placed in a circle at the center of a 60 mm plate and bombarded with approximately 4.7 mg of 1.8 μm tungsten microprojectiles in 1.5 μl aliquots by a

PDS 1000® particle acceleration device. Each plate is bombarded twice through a 150 μm Nytex screen placed about 2 cm above the samples. Freshly bombarded explants are placed in a suspension of transformed *Agrobacterium* for 30 min and then transferred to GBA medium with the cut surfaces down for 3 days with an 18 hr light cycle. Explants are transferred to medium lacking growth regulators but containing drug for selection and grown for 2-5 weeks. After 1-2 weeks more without drug selection leaf samples from green, kanamycin-resistant shoots are assayed for the presence of neophosphotransferase activity. Positive shoots may be grafted to in vitro grown rootstock and transferred to soil.

Within one embodiment of the invention, a genomic DNA sequence containing the END1 or NUC1 promoter region is placed upstream from the *E. coli* β-glucuronidase gene in the plasmid pBI121 (Jefferson et al., *EMBO J. 6*:3901-3907, 1987). in place of the CaMV 35S promoter. This construct is used to transform tobacco leaf discs via *Agrobacterium tumefaciens*-mediated transformation as described herein. Histochemical analysis of tissue sections is performed by adding the substrate 5-bromo-4-chloro-3-indoyl-glucuronide sodium salt in buffer containing potassium ferrocyanide. Plant tissue is incubated in this mixture at 37°C overnight and β-glucuronidase activity is determined colorimetrically (Jefferson, *Plant Molecular Biology Reporter 5*:387-405, 1987).

Any plant that would benefit from expression of a foreign gene is suitable for transformation within the context of this invention. Such plants include barley, corn, oat, wheat, soybean, canola, sunflower and the like. The method of transformation depends, at least in part, on whether the plant is a dicot or monocot.

The following examples are offered by way of illustration, and not by way of limitation.

EXAMPLES

EXAMPLE 1

ISOLATION OF NUC1 AND END1 GENES

cDNA clones of transcripts expressed in the endosperm coenocyte and the nucellus are isolated. Briefly, a cDNA library based on poly(A)-rich RNA of 5 DAP intact ovaries is screened with a positive probe from isolated embryo sacs with appending testa and nucellus cell layers (Fig. 1D) and a negative probe from the remaining pericarp.

Plant Material

Barley (*Hordeum vulgare* L. cv. Bomi) is grown under controlled environmental conditions of 15°C during 16 h light periods and 10°C during 8 hr darkness. At 5 DAP, the barley ovary is 5 mm long (Fig. 1A) and consists of pericarp, testa, nucellus and the endosperm coenocyte at the beginning of the arboreal stage (Fig. 1B). The pericarp has an outer epidermis, multiple layers of parenchyma cells, two layers of photosynthetic cross cells and an innermost loose layer of crushed cells, which permits mechanical separation of the pericarp from the underlying tissues (see below). On the dorsal side of the 5 DAP grain, the nucellus is almost completely degraded, leaving only the nucellar epidermis. On the ventral side, in addition to nucellus itself, the nucellar projection is the dominating tissue. At 8 DAP, when the endosperm has reached the cellular stage, it is enclosed in the nucellar epidermis, which is one cell layer thick (Fig. 1C). Hand-pollinated grains are harvested at appropriate developmental stages, rapidly frozen in liquid nitrogen and stored at -80°C. Individual 5 DAP ovaries are thawed for manual separation of pericarp (negative probe) from the embryo sac with appending testa and nucellus layers (positive probe). After dissection, tissues are rapidly refrozen and stored at -80°C.

Isolation of cDNA clones

A cDNA library of 5 DAP intact ovaries is constructed from poly(A)-enriched RNA by oligo dT-priming and insertion into *Eco*RI-*Xho*I digested λZAPII vector (Clonetech, Palo Alto, CA). The library consists of 1.5 x 10^6 independent clones, with more than 90% of the recombinant plaques containing inserts. About 6000 plaques from this library are absorbed to *E. coli* BB4 cells, plated on LB plates and grown overnight. DNA is transferred to GeneScreen membranes, and the membranes are denatured, neutralized, and affixed to the membranes (*see*, Sambrook et al., *supra*; Ausubel et al., *supra*).

Positive and negative probes for differential screening are prepared from total RNA extracted from embryo sacs with adhering testa and nucellus, and pericarp. Poly(A)-rich RNA is isolated from 70 µg total RNA using oligo$(dT)_{25}$ primer linked to magnetic beads (Dynabeads) according to the manufacturer's instruction (Dynal, Oslo, Norway). Poly(A)-rich RNA is incubated with SuperScript reverse transcriptase (200 units) in the presence of oligo(dT) (50ng), a mixture of dATP, dGTP and dTTP (1.25 mM each), dCTP (0.02 mM), ^{32}P α-dCTP (10 mCi/ml, 3000 Ci/mmol; Amersham, UK) in a final volume of 30 µl buffer (20 mM Tris-HCl buffer, pH 8.4, 50 mM KCl, 2.5 mM $MgCl_2$ 0.1 mg/ml BSA and 10 mM DTT). The reaction is incubated sequentially at 25°C for 10 min, 42°C for 50 min, 90°C for 5 min and 4°C for 10 min followed by 20 min at 37°C in the presence of 10 units of RNAseH.

Differential screening is performed by hybridizing pericarp-specific probe and embryo sac-specific probe to duplicate filters. Prior to hybridization, filters are incubated for 30 min at 42°C in 0.4 M NaOH followed by 30 min at 42°C in 0.1x SSC/0.1% SDS (w/v). Conditions for hybridization are 68 °C in 6x SSC, 0.1% (w/v) SDS, 5x Denhardt's solution, 1 mg/ml salmon sperm DNA and 2 x 10^6 cpm/ml ^{32}P-labeled single-stranded cDNA probe. Filters are washed in 6x SSC and 0.1% (w/v) SDS at 25°C (2x30 min) followed by 68°C(2x30 min). Filters are then exposed to X-ray film at -80 °C for 16 h. Individual plaques that hybridized exclusively with the positive probe are picked and rescreened using the same probes as above.

Confirmed positive phages are excised and converted into Bluescript recombinants using R408 helper phage.

Forty-four clones hybridizing only to the positive probe were selected and grouped into homology classes. Of these classes, six contain two or more clones, whereas the remaining are individual clones that do not hybridize to other positive clones.

Sequence analysis

cDNA inserts are excised from recombinant plasmids by restriction digestion with *Eco*RI and *Xho*I and subcloned into M13 mp18 and mp19. The DNA sequence is determined on both strands using a Sequenase kit (U.S. Biochemical Corporation, OH). Nucleotide and amino acid sequence analyses are performed using the software packages PC/GENE (IntelliGenetics, CA) and GCG (University of Wisconsin).

The sequence of the partial END1 cDNA from the differential screening experiment used as a probe in the experiments reported here is shown in Fig. 4A (SEQ ID NO: 1). The longest open reading frame is presented in SEQ ID NO:3.

The sequence of the NUC1 cDNA isolated in the differential screening experiment is shown in Fig. 4B (SEQ ID NO:2). The open reading frames are presented in SEQ ID NOS:4-6. The function of the NUC1 cDNA remains unknown, as no homologous nucleotide or protein sequence related to NUC1 is present in current databases.

EXAMPLE 2

EXPRESSION OF END1 AND NUC1 TRANSCRIPTS

In cereal grains, such as barley, the endosperm is the main storage organ providing nutrients to the developing and to the germinating embryo. Endosperm ontogeny begins with formation of a diploid embryo by fusion of a sperm

nucleus with the egg cell. Subsequently a triploid endosperm is formed via fusion of a sperm nucleus to the central cell nuclei.

Pre- and post-fertilization structure and development of the megagametophyte in barley is fairly well known (Cass et al., *Can. J. Bot. 63*: 2164, 1985; Engell, *Nord. J. Bot. 9*: 265, 1989; Engell, *Sex Plant Reprod. 7*: 333, 1994). Briefly, barley endosperm development is divided into the syncytial (0-6 days after pollination (DAP), cellularization (6-8 DAP), differentiation (8-21 DAP) and maturation (21-40 DAP) stages. In development, cytokinesis is uncoupled from karyokinesis in the initial stage leading to a large undifferentiated coenocytic stage. After fertilization, three days of rapid nuclear divisions give rise to a population of nuclei evenly distributed in the thin layer of cytoplasm surrounding the vacuole of the central cell. The rate of RNA synthesis increases approximately sevenfold during the syncytial stage (Bosnes et al., *Plant J 2*: 661, 1992), indicating that endosperm development is driven by transcripts synthesized *de novo* from endosperm nuclei. At 5 to 6 DAP, the individual domains undergo marked polarization along an axis perpendicular to the wall of the former central cell. At 6 DAP, the first anticlinal cell walls form in the interzones among nucleocytoplasmic domains. In the following stage, the nuclei enclosed in alveoli of anticlinal walls divide synchronously. Interzonal phragmoplasts then form between separating daughter nuclei, producing the first periclinal wall. Following this round of division, nuclei positioned next to the central vacuole give rise to the starchy endosperm through continued rounds of cell division at random orientation, lasting until 14 DAP (Brown et al., *Plant Cell 6*:1241, 1994).

Within the ovule, the megagametophyte is embedded in the nucellus, believed to be a homologue of the cryptogame megasporangium in which the number of reproductive cells is reduced to only one. At fertilization, the nucellus consists of several cell layers, of which the intermost layer(s) has begun autolysis (Norstog, *Bot. Gazette 35*: 97, 1974). During the following five days of development, all except the outermost nucellar epidermis cell layer disappear, presumably contributing nutrients to the endosperm and the embryo. Over the main ventral vascular strand, nucellus

develops into the nucellar projection, which is responsible for the phloem unloading of assimilates from the sporophyte (Cohrane and Duffus, *Protoplamsa 103*: 361, 1980).

END1 transcripts are differentially expressed in the endosperm coenocyte over the nucellar projection

In transverse sections of 6 DAP grains, an END1 probe hybridizes exclusively to transcripts in the endosperm in the area over the nucellar projection (Fig. 2A). At 6 DAP, the endosperm is at the arboreal stage, and is thickest on the ventral side (see Fig. 1B). No signal is detectable in the dorsal or lateral parts of the coenocyte. As can be seen at higher magnification, the distribution of the END1 transcript is sharply limited to the ventral part of the coenocyte (Fig. 2B). The distribution of silver grains for the END1 probe is bimodal, reflecting the structure of the arboreal endosperm, with a denser accumulation at the top and the bottom than in the central part. No signal is observed in *in situ* hybridization experiments using an END1 sense probe.

In situ hybridization analysis shows that END1 transcripts are present in the cellular endosperm in modified aleurone cells and a few layers of ventral starchy endosperm cells adjacent to the nucellar projection (Fig. 2C and D). In addition, at 8 DAP, a weak signal is detectable also at the dorsal side of the endosperm (Fig. 2C). At 10 DAP, a dorsal signal is not detected.

By Northern blot analysis of poly(A)-rich RNA from whole grains, a single END1 transcript of approximately 920 nucleotides is present in tissue from 5 DAP to 30 DAP, when grain dessication is well underway (Fig. 3, upper panel). The signal is weak during the coenocytic stage, when the endosperm represents a very small fraction of the grain. Subsequently, the signal increases considerably as the endosperm becomes cellular at 8 DAP. After this stage, END1 transcripts are present at a constant steady state level throughout the differentiation and maturation stages, decreasing between 25 and 30 DAP.

Taken together, the data suggest that the modified aleurone cells and the ventral starchy endosperm cells that express END1 transcripts are derived from the part of the endosperm coenocyte which expressed the END1 transcript prior to cellularization. Apart from a weak signal in Northern blots containing poly(A)-rich RNA from 15 and 20 DAP embryos, which may be caused by small adhering pieces of endosperm tissue, no END1 signal is not detectable in any other plant tissue examined, including leaf, stem, root, anthers and immature spikes.

NUC1 transcripts are differentially expressed in the nucellus, nucellar epidermis and in the nucellar projection

In differential screening, the hybridization pattern of NUC1 cDNA is indistinguishable from that of END1, giving a positive signal with probe derived from 5 DAP embryo sacs with adhering testa and nucellus cell layers, and no signal with the pericarp probe. However, in contrast to the endosperm specific END1 clone, NUC1 is detectable by Northern blot analysis of polyA-rich RNA of unfertilized ovules and thereafter from 1 and 6 DAP (Fig. 3, lower panel). The size of the NUC1 transcript is approximately 860 nucleotides. Unless extremely abundant in the central cell and in the young endosperm coenocyte, a transcript of such high relative abundance is most likely to be located in the major sporophytic tissues of the ovule, (*i.e.*, testa or nucellus).

As shown in Fig. 2E, *in situ* hybridization analysis confirms that NUC1 is differentially expressed in the nucellus. No signal is detected in antipodal cells, the central cell, or in the pericarp of unfertilized ovaries. According to Norstog (*Bot. Gazette 35*:97-103, 1974), nucellus degradation is initiated before fertilization, progresses rapidly after fertilization, and is completed around 6 DAP. Thus, the signal between 1 and 6 DAP by in Northern blot analysis comes from transcripts in the degrading nucellus, as shown for the 2 DAP grain in Fig. 2F. Thereafter, when the nucellar parenchyma cells are completely degraded, the NUC1 signal detected in Northern blot analysis is derived from transcripts located in the nucellar epidermis and the nucellar projection (Fig. 2G and H). In addition to nucellar tissues, Northern

blot analysis detects NUC1 transcripts at a low steady-state level in immature and mature anthers, but not in developing or germinating embryos, leaves, stems, immature spikes or roots.

In situ hybridization and microscopy

In situ hybridization is performed as described (Aalen et al., *Plant J. 5*: 385-396, 1994), except that tissue was fixed in 3.7% formaldehyde, 5% acetic acid and 50% ethanol, sections are cut 15-18 μm thick, and autoradiograms are exposed for 6-7 weeks.

Northern blot hybridization

Poly(A)-rich RNA from various grain and vegetative tissues is isolated using magnetic oligo(dT) beads (Dynal, Norway) (Jakobsea et al., *Nucl. Acids Res. 18*: 3669, 1990). Approximately 100 ng of poly(A)-rich RNA from each sample is separated by 1.4% agarose gel electrophoresis and transferred onto nylon membrane filter (Amersham, UK) (*see*, Sambrook et al., *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, 1989). Single-stranded antisense ^{32}P-cDNA probes are generated by amplification of cDNA insert using KS and biotinylated T3 primers (Stratagene, La Jolla, CA) and biotinylated single-stranded cDNA bound to magnetic beads (Dynal, Norway) as template. ^{32}P-labeled probe is synthesized using a random primer labeling kit (Rediprime, Amersham) and α-^{32}P-dCTP (Amersham). Northern filters are hybridized with ^{32}P-labeled probe ($1x10^6$ cpm/ml) at 42°C in the presence of 50% formamide and washed in 2x SSC and 1% SDS at 25°C (2x30 min), followed by washes in 0.2x SSC and 0.1% (w/v) SDS at 68°C (2x30 min). Transcript sizes on the Northern blot are estimated using RNA molecular weight marker I from Boehringer Mannheim on each gel. Filters are exposed for 1 to 3 days. The histone H3 probe used in Fig. 3 to monitor gel loading is isolated from the same cDNA library as used in the differential screening experiment.

EXAMPLE 3

ISOLATION OF PROMOTER REGION OF END1 AND NUC1 GENES

The promoter regions of the END1 and NUC1 genes are isolated from genomic clones. Briefly, a genomic library is constructed in lambda Dash-2 vector (Stratagene, La Jolla, CA) or the EMBL3 vector and packaged *in vitro*. Phage particles are absorbed on *E. coli* and plated on LB plates. Following overnight growth at 37°C, nylon filters are dropped onto chilled plates. Filters are base treated, neutralized, and UV-treated (*see*, Ausubel et al., *supra*). Fragments of END1 and NUC1 are radiolabeled with ^{32}P-dCTP by random priming and synthesis using RTase. Filters are hybridized with labeled probe, washed, and exposed to X-ray film. Hybridizing clones are isolated and propagated. Gene inserts are mapped by restriction enzyme and Southern blotting. Suitable fragments are identified and subcloned. The portion of the insert containing 5' region is isolated by amplification using a primer corresponding to vector sequence and a primer complementary to the 5' region of the END1 and NUC1 gene sequence.

The longest amplified fragments are selected for further study. The fragment is inserted into pBS+ phagemid vector (Stratagene Cloning Systems, La Jolla, CA). Single-stranded phage are recovered after rescue with K07 helper virus. The phagemid DNA is subjected to sequence analysis.

It will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without departing from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Olsen, Odd-Arne
Doan, Danny N.P.
Linnestad, Casper

(ii) TITLE OF INVENTION: ENDOSPERM AND NUCELLUS SPECIFIC GENES, PROMOTERS AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SEED and BERRY LLP
(B) STREET: 6300 Columbia Center, 701 Fifth Avenue
(C) CITY: Seattle
(D) STATE: Washington
(E) COUNTRY: USA
(F) ZIP: 98104

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 30-AUG-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Nottenburg Ph.D., Carol
(B) REGISTRATION NUMBER: 39,317
(C) REFERENCE/DOCKET NUMBER: 750027.402

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (206) 622-4900
(B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 581 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACATTATAAG ATTCTAACAT CAAAGGCATA GAGACGGCTC TAGTGTAGCA GACACAAAAG 60

CCATAGCCAT GGCGAAACTC ATGTGCTTAT GCTTCATCAT CCTCACTATT GCGGTAGTCG 120

TGTCAGCTGG CGGATGCGAC GGTGATCGAG AAGACATGAT CAGGGAGTGT GGTAAGTATC 180

AGAAATTCCC AGCAGAGCCG AAGCTAGCTC CATCAGATGC GTGCTGCGTC GTGTGGCATA 240

AGGCGAACAT CCCATGCCTT TGCGCTGGTG TCACCAAGGA GAAAGAGAAG ATATGGAGCA 300

TGGAGAAGGT TGGCTACGTT GCCAATTTCT GCAAGAAGCC GTTCCCACAT GGCTACAACT 360

GTGGAAGTTA CACATTCCCT CCTCTAGCGT AGTGACTAAT TTATCAGCGG GAGAAGCGTC 420

GGCTTTTACA TTCCATGTTG CTGGCCTCAC AACAACTTTC CTTTGAGATA TGTAATTTGT 480

GAGGTTGTGC TAAAATAAGT TTATGTTAAC CATTATGTGA GAACGTTTGT TATAATAAAG 540

TCTTCTCCTC AAATTGTCTA AAAAAAAAAA AAAAAAAAAA A 581

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 557 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTTCCCTTCG CCAGCACGAG CAAGGAGATG GAGGGGAAGA AGAGAGCGGC GGCCATTGCC 60

GCCCTGTGCA TGGTCCTACT CCTCATGCAG GCAAGGCCAT CCCATCAGCA GTTCTCCGAC 120

TACGCCTGCG AGTGCATCCG GCAGTGCTAC CCCGCGTGCA GGGACAGCAC CCCGCCGTGG 180

CTCTGCAAGA TCAAGTGCGC CGGTAGCTGC CACAACGGCG ACAGGAAGGA CGCACTTACC 240

GCCTGCAGGA TCGCCTGCCT CACGAGCCCC GTCTGCGGCC TGTCGACACC GCCCGTCGCT 300

CCAGGTGATG TTGATCCTTG TACCAGGGAG TGCGACAAGC TGTGGGGTGG CCATGGTCAC 360

GCCAAGGAAC CTTGAAGAGC GACGAGCCGG TCGATCCGGG ATCACATGAT GTCGAATGGC 420

GATGATTGCT GCACCAATCA ATAAAATAAA ATGTATTTCG CTTCGGTGCA TCACTTGTGA 480

GAAAGTTTGT GGGTCTTAGA AGTAAAAGTA AAGATGGTTA TCGACGCATT TGGCGGAAAA 540

AAAAAAAAAA AAAAAAA 557

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

His Gln Arg His Arg Asp Gly Ser Ser Val Ala Asp Thr Lys Ala Ile

```
  1               5                  10                  15
Ala Met Ala Lys Leu Met Cys Leu Cys Phe Ile Ile Leu Thr Ile Ala
            20                  25                  30
Val Val Val Ser Ala Gly Gly Cys Asp Gly Asp Arg Glu Asp Met Ile
        35                  40                  45
Arg Glu Cys Gly Lys Tyr Gln Lys Phe Pro Ala Glu Pro Lys Leu Ala
    50                  55                  60
Pro Ser Asp Ala Cys Cys Val Val Trp His Lys Ala Asn Ile Pro Cys
65                  70                  75                  80
Leu Cys Ala Gly Val Thr Lys Glu Lys Glu Lys Ile Trp Ser Met Glu
                85                  90                  95
Lys Val Gly Tyr Val Ala Asn Phe Cys Lys Lys Pro Phe Pro His Gly
            100                 105                 110
Tyr Asn Cys Gly Ser Tyr Thr Phe Pro Pro Leu Ala
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Leu Pro Phe Ala Ser Thr Ser Lys Glu Met Glu Gly Lys Lys Arg Ala
  1               5                  10                  15
Ala Ala Ile Ala Ala Leu Cys Met Val Leu Leu Leu Met Gln Ala Arg
            20                  25                  30
Pro Ser His Gln Gln Phe Ser Asp Tyr Ala Cys Glu Cys Ile Arg Gln
        35                  40                  45
Cys Tyr Pro Ala Cys Arg Asp Ser Thr Pro Pro Trp Leu Cys Lys Ile
    50                  55                  60
Lys Cys Ala Gly Ser Cys His Asn Gly Asp Arg Lys Asp Ala Leu Thr
65                  70                  75                  80
Ala Cys Arg Ile Ala Cys Leu Thr Ser Pro Val Cys Gly Leu Ser Thr
                85                  90                  95
Pro Pro Val Ala Pro Gly Asp Val Asp Pro Cys Thr Arg Glu Cys Asp
            100                 105                 110
Lys Leu Trp Gly Gly His Gly His Ala Lys Glu Pro
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Phe Pro Ser Pro Ala Arg Ala Arg Arg Trp Arg Gly Arg Arg Glu Arg
 1               5                  10                  15

Arg Pro Leu Pro Pro Cys Ala Trp Ser Tyr Ser Ser Cys Arg Gln Gly
            20                  25                  30

His Pro Ile Ser Ser Ser Pro Thr Thr Pro Ala Ser Ala Ser Gly Ser
        35                  40                  45

Ala Thr Pro Arg Ala Gly Thr Ala Pro Arg Arg Gly Ser Ala Arg Ser
    50                  55                  60

Ser Ala Pro Val Ala Ala Thr Thr Ala Thr Gly Arg Thr His Leu Pro
65                  70                  75                  80

Pro Ala Gly Ser Pro Ala Ser Arg Ala Pro Ser Ala Ala Cys Arg His
                85                  90                  95

Arg Pro Ser Leu Gln Val Met Leu Ile Leu Val Pro Gly Ser Ala Thr
            100                 105                 110

Ser Cys Gly Val Ala Met Val Thr Pro Arg Asn Leu Glu Glu Arg Arg
        115                 120                 125

Ala Gly Arg Ser Gly Ile Thr
    130                 135
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Ser Leu Arg Gln His Glu Gln Gly Asp Gly Gly Glu Glu Glu Ser Gly
 1               5                  10                  15

Gly His Cys Arg Pro Val His Gly Pro Thr Pro His Ala Gly Lys Ala
            20                  25                  30

Ile Pro Ser Ala Val Leu Arg Leu Arg Leu Arg Val His Pro Ala Val
        35                  40                  45

Leu Pro Arg Val Gln Gly Gln His Pro Ala Val Ala Leu Gln Asp Gln
    50                  55                  60

Val Arg Arg
65
```

Claims

We claim:

1. An isolated nucleic acid molecule, comprising an END1 promoter.

2. An isolated nucleic acid molecule, comprising a NUC1 promoter.

3. An isolated nucleic acid molecule, encoding a product of END1.

4. An isolated nucleic acid molecule, encoding a product of NUC1.

5. A vector, comprising a nucleic acid molecule of claims 1 or 2.

6. A vector, comprising a nucleic acid molecule of claims 3 or 4.

7. The vector according to either claims 5 or 6 wherein said vector is an expression vector.

8. The vector of claim 7 wherein the vector is a binary *Agrobacterium tumefaciens* plasmid vector.

9. The vector according to claim 5, further comprising a nucleic sequence encoding a foreign gene operably linked to said promoter.

10. The vector according to either claims 5 or 6, further comprising a selectable marker.

11. The vector according to claim 9 wherein said foreign gene confers disease resistance.

12. A host cell containing a vector according to any one of claims 1-11.

13. The host cell of claim 12 wherein the host cell is a plant cell.

14. The host cell of claim 13 wherein said plant cell is selected from the group consisting of barley, corn, sunflower and soybean.

15. A method of producing a foreign gene product, comprising:

(a) introducing a vector according to claim 9 into a host cell, wherein the vector contains a foreign gene in an expressible form; and

(b) growing the host cell under conditions wherein the foreign gene is expressed.

16. The method of claim 15 wherein said host cell is a plant cell.

17. The method of claim 16 wherein said plant cell is selected from the group consisting of barley, corn, sunflower and soybean.

18. The method of claim 17 wherein transfection is by *Agrobacterium* co-cultivation of bombardment.

19. A method of producing a plant that expresses a foreign gene, comprising:

(a) introducing a vector according to claim 9 into an embryogenic plant cell, wherein said vector contains a foreign gene in an expressible form; and

(b) producing a plant from said embryogenic plant cell, wherein said plant expresses said foreign gene.

20. The method of claim 19 wherein the step of introducing is by *Agrobacterium* co-cultivation or bombardment.

FIG 1A

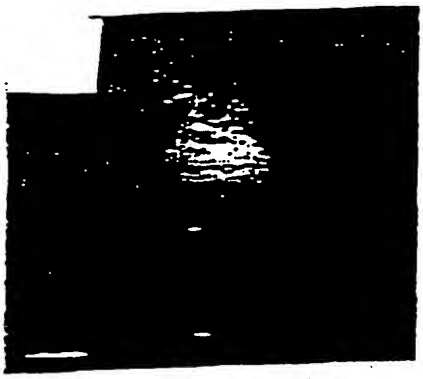
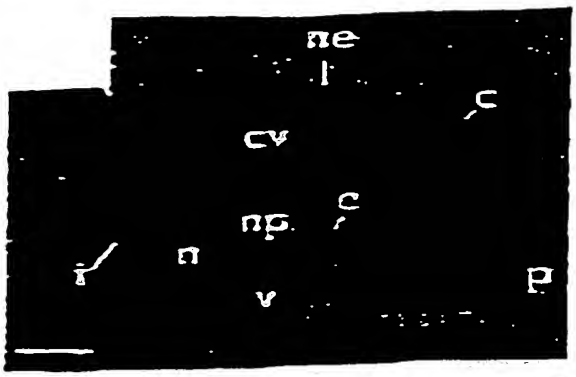


FIG 1B



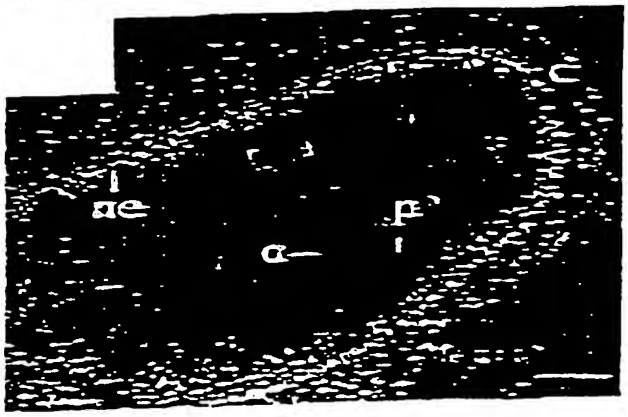


FIG 1C

FIG 1D

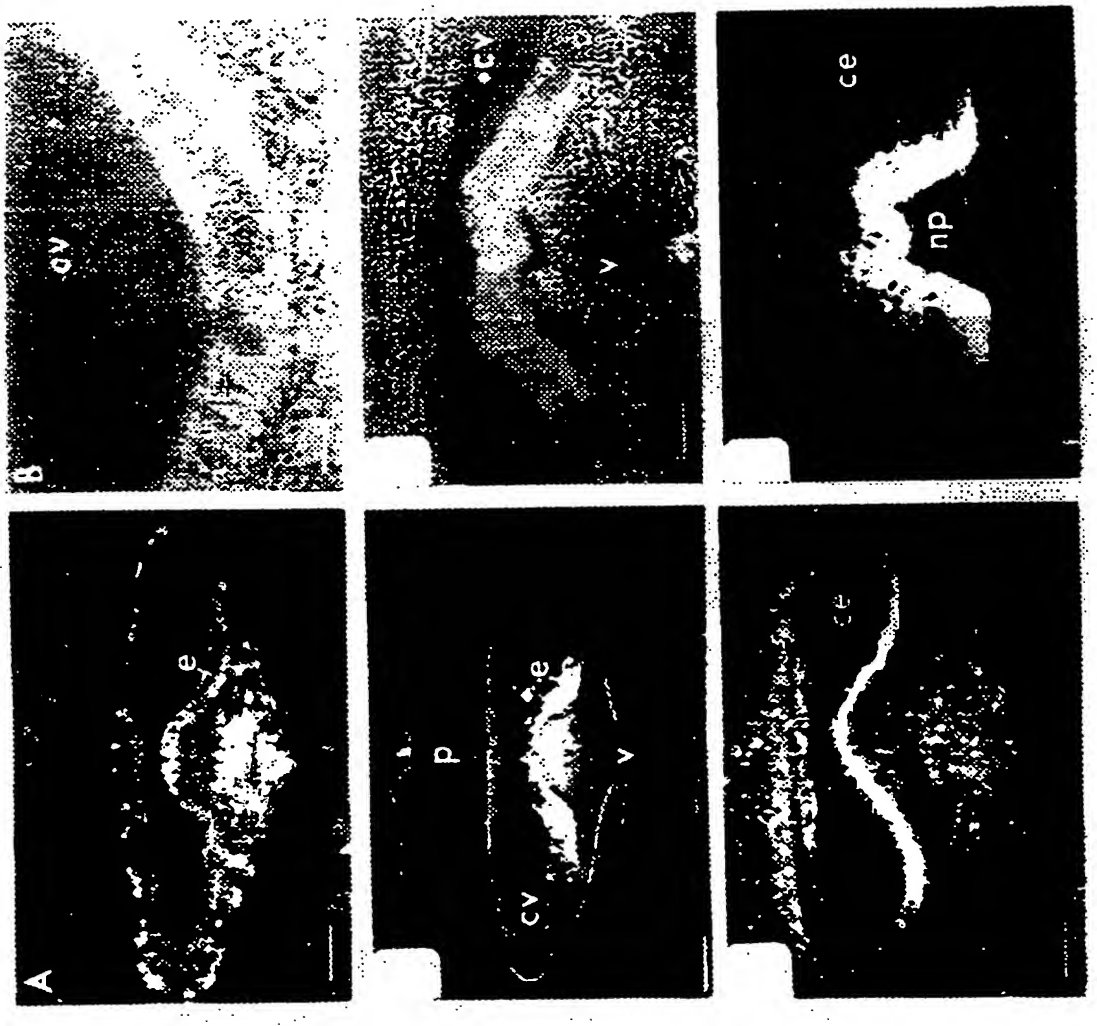


FIG. 2A FIG. 2B

FIG. 2C FIG. 2D

FIG. 2E FIG. 2F

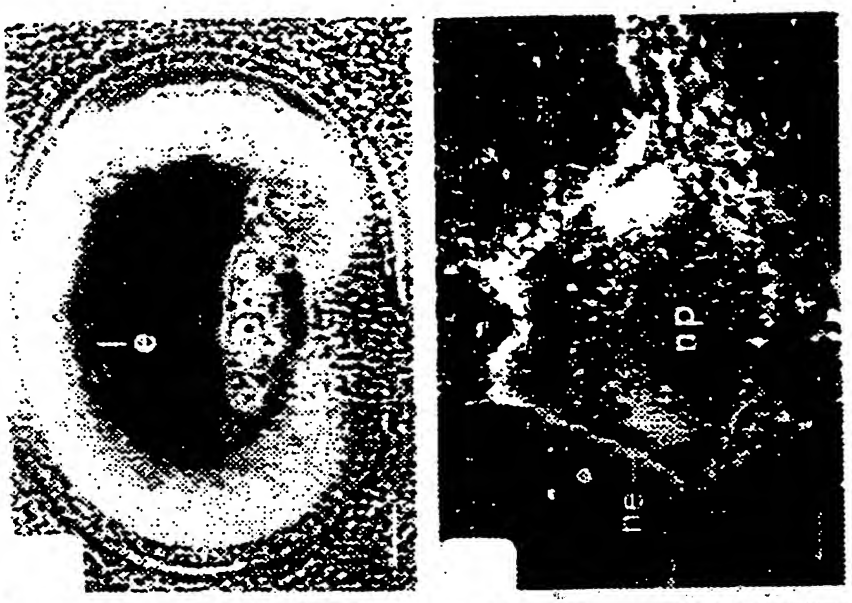


FIG. 2J

FIG. 2H

FIG. 2G

FIG. 2I

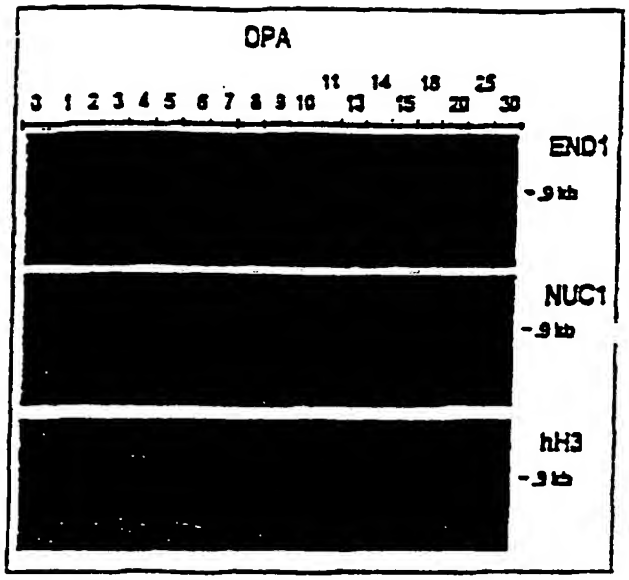
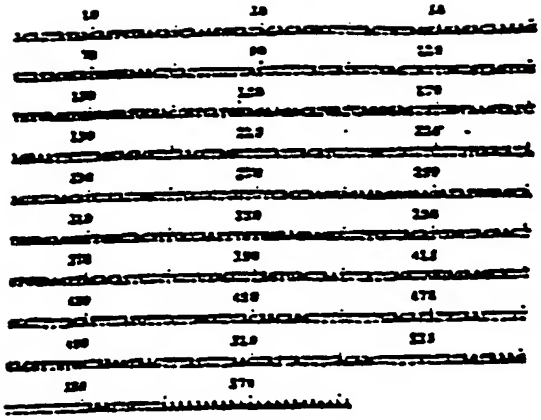


FIG. 3

FIG 4A



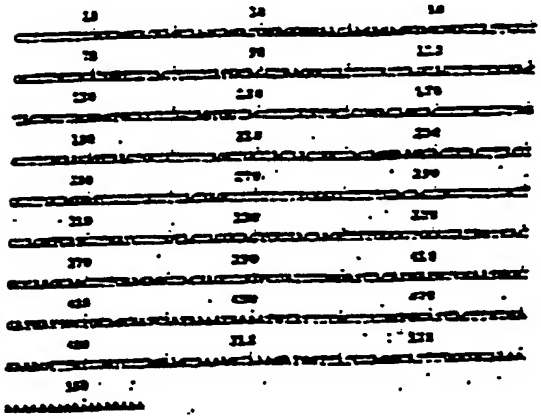


FIG 4B